Virgo computing and storage needs for 2010

November 2009

Virgo collaboration

VSR2/S6 foreseen plan

- Virgo VSR2 possible plan for 2010:

  - about 4 months of stop for upgrading and commissioning in 2010

- LIGO S6 possible plan for 2010:

  - Possible stop starting from October 2010 – still not decided

Virgo storage

- Virgo data to be stored are: raw data, trend and 50Hz data, Reduced Data Set (mainly h_rec) data.
  - Trend, 50Hz and RDS represent a little percentage of the data.
- LIGO data to be stored are h_rec data

- During the runs, we suppose a full data production
- During commissioning, we suppose data production for Science Mode periods which are used for the Astro-watch program.

- Data are stored in the Computing Centers (CCs) of CNAF/Bologna and IN2P3/Lyon.

Virgo storage / data access

- For 2009 we had, until about the mid of the year, VSR1 data stored in disk (Bologna)– HPSS+XrootD cache (Lyon).
  Since VSR2 start (July 7th 2009),
  - ◆ at CNAF VSR1 raw data are stored in the mass storage systems, and VSR2 data on disk.
  - ◆ At IN2P3, VSR1 data have progressively and automatically erased from XrootD cache as they are no more accessed. The XrootD cache will contain mainly VSR2 data, as they are actually analysed.

- For 2010, we suppose to have VSR2 data on disk (Bologna) – HPSS/XrootD cache (Lyon):
  - ◆ The scientific groups have decided to analyze the data as fast as possible, but the procedure is still being tested; there will be also final analysis done by using all the RDS data of the run. Note that **raw data are for example used for candidate followups** through procedures which are still in development. In addition, there is the regular need to **reprocess data quality flags that are defined at some point in the run on earlier data**. This requires to access the whole set of raw data.
  - ◆ For these reasons we need a fast access to VSR2 data.

Virgo storage

- TOTAL VSR2/S6 data:
  - ◆ 75 TB (until October 1st 2009) + 90 TB (October-November 2009) + 245 TB (2010)= **410 TB (<u>NOTA BENE</u>: we use power of 1024 not 1000 to measure sizes)**

- Increase in Virgo storage for 2009:

Period	CNAF disk and Castor [TByte]	IN2P3 XrootD cache and HPSS [TByte]
2009	82 106	44 190

- Situation on October 1st
  - ◆ Disk storage situation at CNAF: **151 TB** free
  - ◆ XrootD cache situation at IN2P3: **184 TB** at disposal in 2009 (140 TB end 2008 + 44 TB in 2009)

- Probable situation at the end of 2009:
  - ◆ CNAF 60 TB free.
  - ◆ IN2P3 : the cache disk might not be fully used (19 TB left)

Virgo storage

- Maximum increase in Virgo storage needs for 2010:

Period	CNAF disk and Castor [TByte]	IN2P3 XrootD cache and HPSS [TByte]
2010	185 20	226 276

- ◆ CNAF disk: 245 TB (8 months Virgo) – 60 TB free space remaining
- ◆ IN2P3 XrootD cache: 245 TB – 19 TB free space remaining

- Astrowatch data 30 TB: to be stored in CASTOR and probably not accessed often
- The increase with respect to the last estimation is due to the increase in the data acquisition rate (about 11Mb/s for VSR2) and the plan to have a commissioning period for 2010 of 4 instead of 6 months as previously stated.

Virgo storage / data access

- The situation is different in CNAF/Bologna:
  - ◆ CNAF: CASTOR is the mass storage system on tape, but data are not accessed from CASTOR directly
  - ◆ IN2P3: data are stored permanently in HPSS and are staged in the XrootD cache according to users' access.

- For IN2P3 we could estimate a use of 75% of the total volume of VSR2/S6 data, thus about 308 TB.
  - ◆ The **increase** for XrootD cache in 2010 could then be limited to:

    - » 308 TB – 184 TB = **124 TB**
    - » **This request could be revised if we encounter problems in accessing data in Lyon in 2010 (the factor of 75% is a rough estimate based on past years experience).**

Virgo computing - 2009

- Virgo Data Analysis is mainly done in the CCs where all the data are at disposal.
- 2009 Computing situation:
  - Until October 1st, 90.000 kSPCint2k.day where used at CNAF, 77.000 at IN2P3.
  - Taking into account that some searches have been delayed, we can lower the estimate previously done:

	CNAF/Bologna [kSPECINT2000.day]	IN2P2/Lyon [kSPECINT2000.day]
Continuous signals	200,000	0,0
Burst sources	5,000	100,000
Stochastic Background	5,000	0
Coalescing Binaries	70,000	50,000
Detector Characterization	5,000	5,000
Total	285,000	155,000

- A CNAF/Bologna, in case of overloading, only 10% of the farm computing energy is dedicated to Groups other than Group 1 experiments, so that mainly at the end of each year we keep experiencing problems in the Virgo jobs execution.

Virgo computing - 2010

- For 2010, we hope to run part of the burst and CBC searches in Bologna and Lyon. We will need then more computing energy:

	CNAF/Bologna [kSPECINT2000.day]	IN2P2/Lyon [kSPECINT2000.day]
Continuous signals	200,000	0,0
Burst sources	10,000	200,000
Stochastic Background	5,000	0
Coalescing Binaries	100,000	100,000
Detector Characterization	5,000	5,000
Total	320,000	305,000